

卒業論文 2016 年度（平成 28 年度）

自動打ちシステムの実戦による  
麻雀戦略の検証

慶應義塾大学 環境情報学部

細田 航星

徳田・村井・楠本・中村・高汐・バンミーター・植原・三次・中澤・武田  
合同研究プロジェクト

2017 年 1 月

卒業論文 2016 年度（平成 28 年度）

## 自動打ちシステムの実戦による 麻雀戦略の検証

### 論文要旨

近年、2 人零和確定完全情報ゲームであるチェス、オセロ、将棋といったゲームでは人間のトッププレイヤーと同等かそれ以上の実力を持つコンピュータプレイヤーが提案されている。一方で多人数不完全情報ゲームである麻雀では未だそれに及ばない状態である。これは従来のゲーム木の探索手法を適用するのが難しいためである。本論文では、探索空間が大きくなることでそのまま適用できなかったモンテカルロ法について、数理的な評価と適用する部分を限定することで、適用した。また、オリについては牌譜の統計的な分析を行うことで、読みの精度を上げるアプローチを取った。評価として、人間のプレイヤーや先行研究 AI との比較を行った。解析の結果、本手法で構築した AI は平均レベルのプレイヤーに匹敵する実力を出すことができたことがわかった。

### キーワード

慶應義塾大学 環境情報学部

細田 航星

# Abstract Of Bachelor's Thesis    Academic Year 2016

## Summary

Eigo ga dekinai node Roma-ji de soreppoi hunniki wo daseruto iina.

Murippoi desu ne.

Write down your abstract here. Write down your abstract here. Write down your abstract here. Write down your abstract here. Write down your abstract here. Write down your abstract here.

Write down your abstract here. Write down your abstract here. Write down your abstract here. Write down your abstract here. Write down your abstract here. Write down your abstract here. Write down your abstract here. Write down your abstract here. Write down your abstract here. Write down your abstract here.

Write down your abstract here. Write down your abstract here.

## Keywords

Bachelor of Arts in Environmental Information  
Keio University

Wataru Hosoda

# 目 次

第 1 章	序論	1
1.1	本論文の背景 . . . . .	1
1.2	本論文が着目する課題 . . . . .	1
1.3	本論文の目的 . . . . .	2
1.4	本論文の構成 . . . . .	2
第 2 章	結論	3
2.1	本研究のまとめ . . . . .	3
2.2	今後の課題と展望 . . . . .	3
	謝辞	4
	参考文献	5

## 图 目 次

## 表 目 次

# 第1章 序論

## 1.1 本論文の背景

近年、2人零和確定完全情報ゲームであるチェス、オセロ、将棋といったゲームではAIの研究がめざましく、人間のトッププレイヤーと同等以上の成績を記録している。一方で多人数不完全情報ゲームである麻雀では未だそれに及ばない状態である。これは従来のゲーム木の探索手法を適用するのが難しいためである。多人数ゲームでは状況によってプレイヤーの行動の目的が違うといった難しさがある。この難しさを解消するために、先行研究 [4][3] でも多数行われている通り、相手を考慮しない1人麻雀プレイヤーに対して和了率を求める評価を採用する。麻雀を行う時の人間の思考は1人麻雀をもとにしているため、1人麻雀から4人麻雀への拡張は容易だと考えられる。

## 1.2 本論文が着目する課題

1人麻雀においてある局面からゲーム木を展開しようとした場合、次にどの種類牌をツモるかはわからないため、ランダムに決定しノードを展開していくことになる。これを繰り返し行くと探索空間が大きくなりすぎるため、手の選択が難しくなる。これを対処するために、あらゆる手法が提案されている。このような場合に木を展開せずに有望な手を選択する手法として、Upper Confidence Bound (UCB) [1] がある。UCBではある局面から考えられる全ての手に対して、よい結果を返しそうな手を重視しつつ、何度もプレイアウトと呼ばれるランダムシミュレーションを行うことで最善の手を決定する手法である。UCBは局面ごとにこの探索を実行するが、各局面を完全に別の局面として扱うため他の局面の探索において得た情報を利用することができない。この問題を解決するために Linear UCB (LinUCB) [2] という手法が提案されている。これは、局面を特徴で表すことで、対象とする局面が異なっても、それまで対象とした局面の情報を利用できるといったものである。これらは麻雀について適用した例が報告されている [4] が、いずれの方法も平均プレイヤーに満たない成績となった。

本研究では、モンテカルロ法の問題点である、麻雀に適用すると探索空間が大きくなりすぎてうまく適用出来ない問題に対して、探索空間を小さくするために、2つの解決策を行う。一つはモンテカルロ法のシミュレーションを直接牌姿に当てはめるのではなく、牌姿ごとの局面で数理的に比較できる内容を数理計算を用い削減する方法である。もう一つは、三向聴以下に対してはシャンテン数を下げるように打ち、モンテカルロ法を用いる部分を限定する方法である。麻雀の牌理に関しては、上がりに近づく選択ほど重要である [5] ため、聴牌に近い打牌選択においてモンテカルロ法を適用することで、成績に影響を大きく与える部分でモンテカルロ法の効果が発揮できることが期待される。

### 1.3 本論文の目的

本研究では和了率を理論値計算によって評価することと和了に近い段階へのモンテカルロ法の部分適用によって、麻雀に対してモンテカルロ法をそのままうまく適用できない問題に対して解決策を提案する。多人数ゲームでは状況によってプレイヤーの行動の目的が違ふといった難しさがある。この難しさを解消するために、先行研究 [1,2] でも多数行われている通り、相手を考慮しない 1 人麻雀プレイヤーに対して和了率を求める評価を採用する。従来のモンテカルロ法を麻雀に適用した例では平均プレイヤーレベルの実力が出ていないため、本手法によってそれらを上回ることが期待される。

### 1.4 本論文の構成



## 第2章 結論

### 2.1 本研究のまとめ

本稿では、牌譜の局面からの教師あり学習や異なる探索の結果の共有ができる LinUCB を 1 人麻雀に適用することを提案した。LinUCB と、比較手法として UCB1 と事前学習を利用する手法を 1 人麻雀に適用し、あがった局数を比較することで評価を行った。その結果いずれの手法も人間のプレイヤーには及ばず、3 つの手法では UCB1 が最も良い あがり率を出す結果となった。しかし、UCB1 と事前学習のみを用いる手法の性能が近かったことから、事前学習が不完全情報ゲームに対してある程度有効であるということが分かった。また、Plain LinUCB, 事前学習+LinUCB の結果から特徴量の設計が不適切であったことが考えられる。今後は特徴量の見直して実験を行い、LinUCB の性能評価を再度行いたい。本実験では、LinUCB と事前学習を利用する手法を分けて評価を行ったが、LinUCB で成果が出るようになれば、これらを組み合わせることで更に性能が向上することが期待されるのでこれについても評価を行いたい。また、UCB をモンテカルロ木探索に適用した UCT [8] が成果をあげていることから LinUCB をモンテカルロ木探索へ適用することも検討している。参考文献

### 2.2 今後の課題と展望

# 謝辞

本研究を進めるにあたり、ご指導いただきました慶應義塾大学 環境情報学部教授 村井純博士、同学部教授 中村修博士、同学部准教授 Rodney D. Van Meter III 博士、同学部准教授 植原啓介博士、同学部准教授 中澤仁博士、SFC 研究所 上席所員 (訪問) 齊藤賢爾博士に感謝致します。

研究について日頃からご指導頂きました政策・メディア研究科博士課程 松谷健史氏、政策・メディア研究科特任助教 空閑洋平氏、XXXXXX 徳差雄太氏に感謝致します。研究室に所属したばかりの頃から本研究に至るまで、特定の分野にこだわらない広い視点で何年生の時であっても妥協のない姿勢で向かい合い、絶えず多くのご指導をいただきました。本研究を卒業論文としてまとめることができたのも両氏のおかげです。重ねて感謝申し上げます。

本研究の評価に必要な伝送装置の助言、機材を運搬していただいた一般社団法人 Mozilla Japan 工藤紀篤博士に感謝いたします。評価に必要な伝送装置を借用させていただいた慶應義塾大学デジタルメディア・コンテンツ統合研究センターの皆様には感謝いたします。長期の間、開発、実験用に 4K カメラなどの機器を借用させていただいた慶應義塾大学湘南藤沢メディアセンターマルチメディアサービスの皆様には感謝いたします。

研究室を通じた生活の中で多くの示唆を与えてくれた XXXX 氏、および Arch 研究グループの皆様には感謝します。また、徳田・村井・楠本・中村・高汐・バンミーター・植原・三次・中澤・武田 合同研究プロジェクトの皆様には感謝致します。

最後に、私の研究を支えてくれた両親をはじめとする親族、多くの友人・知人に感謝し、謝辞と致します。

## 参考文献

- [1] Auer, p., cesa-bianchi, n. and fischer, p.: Finite-time analysis of the multiarmed bandit problem, machine learning, vol. 47, no. 2-3, pp. 235256 (2002).
- [2] Li, l., chu, w., langford, j. and schapire, r. e.: A contextual-bandit approach to personalized news article recommendation, proceedings of the 19th international conference on world wide web, acm, pp. 661670 (2010).
- [3] 水上 直紀ほか. 降りるべき局面の認識による 1 人麻雀プレイヤーの 4 人麻雀への適用. 第 18 回 ゲームプログラミングワークショップ (GPW2013) , Vol. 31, No. 3, pp. 1–7, 2013.
- [4] 中張 遼太郎ほか. Linucb の 1 人麻雀への適用. ゲームプログラミングワークショップ, 2013.
- [5] ネマタ. テンパイの技術. 勝つための現代麻雀技術論, 2014.